

# Modelling the broadband propagation of marine mammal echolocation clicks for click-based population density estimates

Alexander M. von Benda-Beckmann<sup>a)</sup>

*The Netherlands Organisation for Applied Scientific Research (TNO)—Technical Sciences, Sonar Group, 2597 AK, The Hague, The Netherlands*

Len Thomas

*Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, Fife KY16 9LZ Scotland, United Kingdom*

Peter L. Tyack

*Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, St. Andrews, Fife KY16 8LB Scotland, United Kingdom*

Michael A. Ainslie<sup>b)</sup>

*The Netherlands Organisation for Applied Scientific Research (TNO)—Technical Sciences, Sonar Group, 2597 AK, The Hague, The Netherlands*

(Received 9 August 2017; revised 30 December 2017; accepted 21 January 2018; published online 14 February 2018)

Passive acoustic monitoring with widely-dispersed hydrophones has been suggested as a cost-effective method to monitor population densities of echolocating marine mammals. This requires an estimate of the area around each receiver over which vocalizations are detected—the “effective detection area” (EDA). In the absence of auxiliary measurements enabling estimation of the EDA, it can be modelled instead. Common simplifying model assumptions include approximating the spectrum of clicks by flat energy spectra, and neglecting the frequency-dependence of sound absorption within the click bandwidth (narrowband assumption), rendering the problem amenable to solution using the sonar equation. Here, it is investigated how these approximations affect the estimated EDA and their potential for biasing the estimated density. EDA was estimated using the passive sonar equation, and by applying detectors to simulated clicks injected into measurements of background noise. By comparing model predictions made using these two approaches for different spectral energy distributions of echolocation clicks, but identical click source energy level and detector settings, EDA differed by up to a factor of 2 for Blainville’s beaked whales. Both methods predicted relative density bias due to narrowband assumptions ranged from 5% to more than 100%, depending on the species, detector settings, and noise conditions. © 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5023220>

[KL]

Pages: 954–967

## I. INTRODUCTION

Passive acoustic monitoring (PAM) has potential for cost-effective surveillance of the population density of many marine mammal species (as reviewed in [Marques \*et al.\*, 2013](#)). Converting counts of detected vocalizations into estimates of spatial population density requires, among other things, an estimate of the area around each PAM receiver over which vocalizations are detected—the “effective detection area” or EDA. This EDA is typically obtained by estimating the probability of detecting a vocalization as a function of distance (the “detection function”) and then averaging it in an appropriate way (described in detail later). Several methods exist to estimate the detection function ([Marques \*et al.\*, 2013](#)); the most reliable require measurements of auxiliary information such as the range to detected vocalizations. Distances to vocalizing animals can sometimes be estimated directly through passive

sonar localisation ([Watkins and Schevill, 1972](#); [McDonald and Fox, 1999](#); [Marques \*et al.\*, 2011](#); [von Benda-Beckmann \*et al.\*, 2013](#); [Zimmer, 2013](#); [Gassmann \*et al.\*, 2015](#)), or when ground-truth information is available, for instance, through acoustic tags attached to vocalizing animals ([Marques \*et al.\*, 2009](#); [Ward \*et al.\*, 2011](#)). However, options are limited when single sensors are deployed in a wide-baseline array such that they operate independently. In many cases (e.g., [Küsel \*et al.\*, 2011](#); [Harris, 2012](#); [Hildebrand \*et al.\*, 2015](#); [Frasier \*et al.\*, 2016](#)), one must rely on an acoustical modeling approach, where one assumes a source level (or distribution of source levels), models signal propagation at the specified range, and then determines probability of detection using the passive sonar equation. In such cases, the reliability of the density estimation depends on the fidelity of the modeling of detection.

Different approaches have been adopted to model detection functions for echolocating marine mammals, focusing on different aspects of the detection problem, such as detecting single clicks ([Marques \*et al.\*, 2009](#)), detecting multiple clicks ([Zimmer \*et al.\*, 2008](#); [von Benda-Beckmann \*et al.\*, 2010](#); [Küsel \*et al.\*, 2011](#)), or assessing how varying noise

<sup>a)</sup>Electronic mail: [sander.vonbendabeckmann@tno.nl](mailto:sander.vonbendabeckmann@tno.nl)

<sup>b)</sup>Present address: JASCO Applied Sciences (Deutschland) GmbH, Mergenthalerallee 15-21, 65760 Eschborn, Germany.

affects the detection function (Ward *et al.*, 2011). Model predictions for detection functions are obtained either using the passive sonar equation to model the detection process (Zimmer *et al.*, 2008; Kimura *et al.*, 2010; Küsel *et al.*, 2011; Harris, 2012; Ainslie, 2013; Hildebrand *et al.*, 2015; Frasier *et al.*, 2016), or applying detectors to detect modeled vocalizations (clicks or calls) injected into recorded noise (von Benda-Beckmann *et al.*, 2010; Helble *et al.*, 2013; Küsel *et al.*, 2016).

In order to predict the detectability of broadband echolocation clicks, assumptions have to be made about propagation loss and the spectral energy distribution of the source waveform. Different approaches have been adopted to characterize the click source level in the passive sonar equation, for instance, by using the source factor (Ainslie 2010; ISO, 2017) measured over the duration of the click (Zimmer *et al.*, 2008; von Benda-Beckmann *et al.*, 2010; Küsel *et al.*, 2011; Ward *et al.*, 2011), or by its peak-to-peak sound pressure (Au, 1993; Hildebrand *et al.*, 2015; Küsel *et al.*, 2016). Alternatively, the spectral energy distribution of a click's source waveform can be approximated by a top-hat function, within the effective bandwidth of the click (Ainslie, 2013). These procedures often implicitly or explicitly assume that the energy is uniformly distributed within a predefined bandwidth (the detection bandwidth, or the click frequency band), which is a simplification, because in reality echolocation clicks have non-uniform and variable spectral energy distributions (Møhl *et al.*, 2003; Wahlberg *et al.*, 2011; Shaffer *et al.*, 2013).

Another common simplification is that the propagation factor (combination of spreading loss and absorption) is independent of frequency across the frequency band of the click (Zimmer *et al.*, 2008; Küsel *et al.*, 2011; Ward *et al.*, 2011; Hildebrand *et al.*, 2015; Frasier *et al.*, 2016). Previous work has shown that the bandwidth of an echolocation click has an important effect on the propagation loss (PL) (Ainslie, 2010) and therefore on the detection probability (von Benda-Beckmann *et al.*, 2010; Ainslie, 2013) and that potentially large errors in detection range can arise from such narrowband methods (Ainslie, 2013). The effect on detection probability remained unclear because neither von Benda-Beckmann *et al.* (2010) nor Ainslie (2013) considered the effect of click bandwidth on the figure-of-merit, a quantity that depends on processing bandwidth, signal and noise statistics and choice of detector.

The present paper investigates the extent to which the broadband nature of echolocation clicks needs to be accounted for when modeling their detectability, by comparing different model approaches: (1) by simulating the detection function using the passive sonar equation with different approximations of the echolocation click or (2) by injecting recorded echolocation clicks into real noise. The second method has the following advantages: it naturally accounts for the spectral shape of the click, it more easily accounts for frequency dependent absorption, and it can empirically measure the detection functions for detection processes that are not easily modeled in the passive sonar equation.

To assess the potential differences in predicted detection functions for clicks, recorded on-axis echolocation clicks were selected and propagated (similar to von Benda-

Beckmann *et al.*, 2010; Helble *et al.*, 2013; Küsel *et al.*, 2016) and added to recorded background noise in the same frequency band as the click. The present study considers the detection of Blainville's beaked whale (*Mesoplodon densirostris*), as this and other beaked whale species have been the focus of different studies to estimate animal density using passive acoustics (Marques *et al.*, 2009; Küsel *et al.*, 2011; Hildebrand *et al.*, 2015). The same methodology is applied also to sperm whale (*Physeter macrocephalus*) clicks, whose high source level ensures long range detection and hence amplifies differential absorption effects.

The detection function for a particular signal depends on the detection process adopted, and what false-alarm rate (FAR) is deemed acceptable. A wide range of detectors is commonly used for echolocation clicks, such as different energy-based detectors (Gillespie and Leaper, 1996; Mellinger, 2001; Yack *et al.*, 2010; Klinck and Mellinger, 2011), peak detectors (Hildebrand *et al.*, 2015), a Teager-Kaiser detector (Kandia and Stylianou, 2005), a Page test detector (van IJsselmuide and Beerens, 2004) or a matched filter (Ward *et al.*, 2008). In the present study, two different click detectors are implemented and applied to obtain detection functions, in order to assess the influence of the detection process on the estimated relative bias in detection range and population density. First, a simple sub-band energy detector is applied, facilitating a relatively straightforward comparison with the estimates from Ainslie (2013) using the passive sonar equation. Second, the power-law Page test detector (von Benda-Beckmann *et al.*, 2010) is adopted as an example of a more complex detector.

The results obtained in this way are compared with those obtained by approximating the propagation loss using the narrowband approximation, in which all click energy is assumed to propagate at the click center frequency. The findings of typical detection thresholds are used in the broadband passive sonar approximation of Ainslie (2013), to estimate the range of relative biases due to narrowband approximation of absorption loss for echolocating species other than Blainville's beaked whales and sperm whales. Finally, different model approaches for the spectral energy distribution of the echolocation clicks, and their effect on the effective detection areas are compared.

## II. METHODS

### A. Determining the effective detection area

#### 1. Relationship between the detection function and effective detection area

Following Marques *et al.* (2013), the animal density  $D$  from  $n$  detected cues (in this case echolocation clicks), is estimated using the equation

$$\hat{D} = \frac{n}{\hat{A}_{d,\text{eff}}} \times \hat{K}, \quad (1)$$

where  $A_{d,\text{eff}}$  is the effective detection area (EDA),  $K$  combines other multipliers used to convert the number of cues into animals, and a circumflex indicates a quantity that is estimated (i.e., not known). The EDA is defined as

$$A_{d,\text{eff}} = 2\pi \int_{r=0}^{\infty} r g(r) dr, \quad (2)$$

where  $r$  is range and  $g(r)$  is the detection function, i.e., the probability of detection at range  $r$ . It can be interpreted as the area of a circle around the detector such that the number of cues detected outside the circle is equal to the number missed within it (see Buckland *et al.*, 2001, p. 54, for details). Note that, in practice, detection probability becomes zero at some range,  $w$ , so the upper limit of integration in Eq. (2) can be truncated at  $w$ . Marques *et al.* (2009) give an equivalent formulation, but instead of  $\hat{A}_{d,\text{eff}}$  in Eq. (1) they wrote  $\pi w^2 \hat{P}$ , where  $P$  is the average probability of detecting a cue within distance  $w$  of the sensor and  $P = 2 \int_{r=0}^w r g(r) dr / w^2$ .

In this study we consider situations where the detection function  $g(r)$  is estimated using acoustic models, and assess how the assumptions when modelling broadband propagation of an echolocation click affects the estimated animal density. For this purpose, we introduce the term *relative bias* in the estimated animal density, which quantifies the bias introduced by using a narrowband approximation of the propagation when modeling the detection function,

$$\begin{aligned} \text{relative density bias} &= \frac{E(\hat{D}_{\text{NB}} - E(\hat{D}_{\text{BB}}))}{E(\hat{D}_{\text{BB}})} \\ &= \frac{E(\hat{A}_{d,\text{eff,NB}} - E(\hat{A}_{d,\text{eff,BB}}))}{E(\hat{A}_{d,\text{eff,NB}})}, \end{aligned} \quad (3)$$

where subscript BB denotes broadband and NB denotes narrowband, and  $E(\hat{x})$  denotes the expected value (i.e., average) of the estimate  $\hat{x}$ . It is important to note that the estimated biases are relative, not absolute—we have compared the error in density due to the narrowband assumption relative to the broadband method. While both are approximations, the broadband propagation approach with the recorded waveform is considered to more closely approximate the real-world situation.

The effective detection radius  $R_{d,\text{eff}} = \sqrt{A_{d,\text{eff}}/\pi}$  is the range at which as many detections would be detected outside this range as would be missed inside it. This radius was also computed, assuming a uniform spatial and temporal distribution. Note that  $R_{d,\text{eff}}$  should not be confused with the related concept of the detection range,  $r_{50\%}$  (the distance at which  $g = 0.5$  for a specified false alarm probability) which, while commonly estimated from the passive sonar equation (Zimmer *et al.*, 2008; von Benda-Beckmann *et al.*, 2010; Ainslie, 2013), is of little practical use in the context of animal density estimation.

## B. Simulating echolocation clicks in recorded background noise

Synthetic datasets were generated by propagating a recorded waveform of an echolocation click and injecting the propagated waveform into segments of measured background. Then two different detectors, a simple sub-band energy detector and a power law Page test detector, were

applied to this simulated dataset and the detection performance was quantified. Thereby, detection functions  $g(r)$  for realistic detector settings (and corresponding false-alarm rates) were constructed.

A synthetic calibrated source waveform (ISO, 2017),  $s(t)$ , for each species was constructed by taking a measured source time series,  $x_{\text{src}}(t)$ , and scaling according to the energy source level ESL (in dB re  $1 \mu\text{Pa}^2 \text{ m}^2 \text{ s}$ ), obtained from the literature for that species,

$$s(t) = x_{\text{src}}(t) \frac{\sqrt{S_E}}{\sqrt{\int_0^{\infty} x_{\text{src}}(t)^2 dt}}, \quad (4)$$

where  $S_E$  is the energy source factor (ISO, 2017), related to the energy source level (ESL) via

$$\text{ESL} = 10 \log_{10} \frac{S_E}{1 \mu\text{Pa}^2 \text{ m}^2 \text{ s}} \text{ dB}. \quad (5)$$

On-axis source signals were adopted in this study for the two study species (Fig. 1): Blainville's beaked whale (*Mesoplodon densirostris*), and sperm whale (*Physeter macrocephalus*). The click energy source level was assumed to be normally distributed around the mean energy source level (Küsel *et al.*, 2011). The estimated mean and standard deviation of the click energy source levels for each species were adopted from the literature (Table I).

For propagating the source waveform to a receiver at a distance  $R$  from the source, a simple spherical spreading law and an absorption term,  $\alpha(f)$ , to account for the frequency dependent attenuation of sound, were adopted. This simplified model was adopted to highlight effects from other assumptions, but more complicated models may be required for real-life applications of detection functions (see Sec. IV). The transfer function for spherical spreading,  $H(R, f)$ , was computed as

$$H(R, f) = \frac{1}{R} e^{-\alpha(f)R} e^{-2\pi i f (R/c_s)}, \quad (6)$$

with the frequency  $f$ , the sound speed  $c_s$ , and  $i$  the imaginary unit. A fixed sound speed of 1500 m/s was adopted throughout this study. The transfer function was used to multiply the complex frequency spectrum  $P(R, f)$  of the source waveform, given by  $P(R, f) = S(f)H(R, f)$ , to achieve the spectrum of the signal at the receiver. The received signal sound pressure,  $p(R, t)$ , was then obtained by inverse-FFT of  $P(R, f)$ .

Two cases were adopted for modeling propagation loss: a broad-band case, where the absorption was modeled to be frequency dependent over the bandwidth of the signal; and a narrow-band case, where the transfer function  $H_{\text{NB}}(R, f)$  was approximated by computing it at the central frequency of the source spectrum,  $f_c$ , and assumed to apply to the entire frequency band. Figure 2 illustrates modeled received waveform for the Blainville's beaked whale echolocation click propagated to different distances (1000, 3000, and 5000 m) using the broadband and narrowband assumptions for absorption.

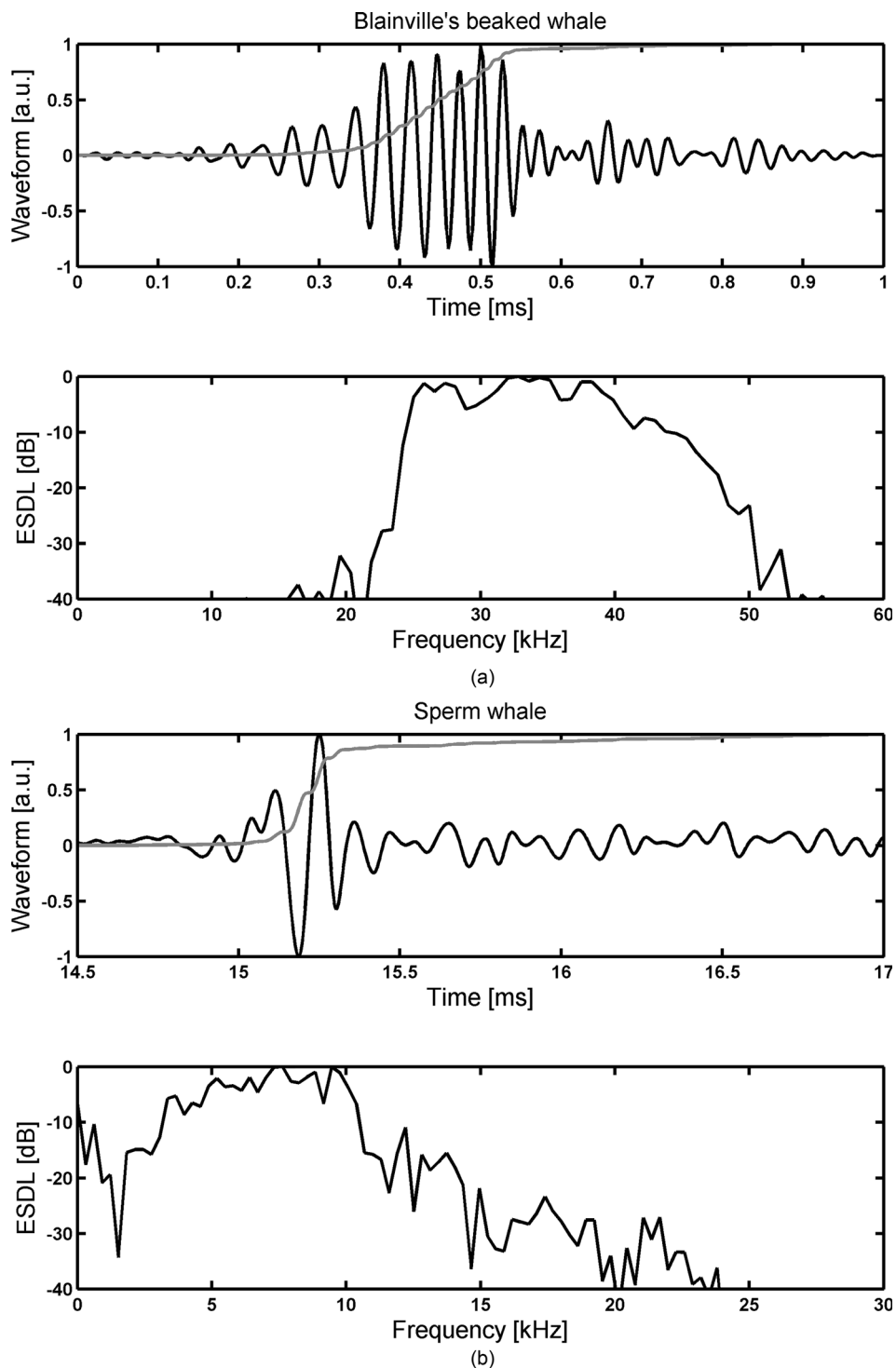


FIG. 1. On-axis source waveforms (black, top panels), and energy spectral density level (ESDL) (black, bottom panels) of an echolocation click used as source waveforms in this study. The cumulative energy is indicated in gray superimposed on top of the pressure waveforms. The top panel shows an example of an on-axis sperm whale click (from Zimmer, 2011), the bottom panel shows an example of an on-axis Blainville's beaked whale click (from von Benda-Beckmann *et al.*, 2010). Waveforms and ESDL were normalized to the maximum values.

TABLE I. Source parameters for species-specific echolocation clicks, the mean energy source level and standard deviation, and detector frequency range  $[f_{\min}, f_{\max}]$ , adopted to estimate detection functions in simulated data. The last column lists the mean energy noise level ENL within these frequency bands as measured in the AUTECH09 dataset, and for typical wind noise and thermal-limited background spectral noise levels for a sea state of SS = 2, based on Wenz (1962).

Species	Mean ESL [dB re $1 \mu\text{Pa}^2 \text{m}^2 \text{s}$ ]	STD (ESL) [dB]	$f_{\min}, f_{\max}$ [kHz]	Reference	ENL—AUTECH09 [dB re $1 \mu\text{Pa}^2 \text{s}$ ]	ENL—Wenz (1962) for Sea State 2 [dB re $1 \mu\text{Pa}^2 \text{s}$ ]
<i>Mesoplodon densirostris</i> Blainville's beaked whale	164	1.8	24, 48	Küsel <i>et al.</i> (2011)	70.8	40.8
<i>Physeter macrocephalus</i> Sperm whale	190	2.5	2, 20	Möhl <i>et al.</i> (2003)	83.3	50.0



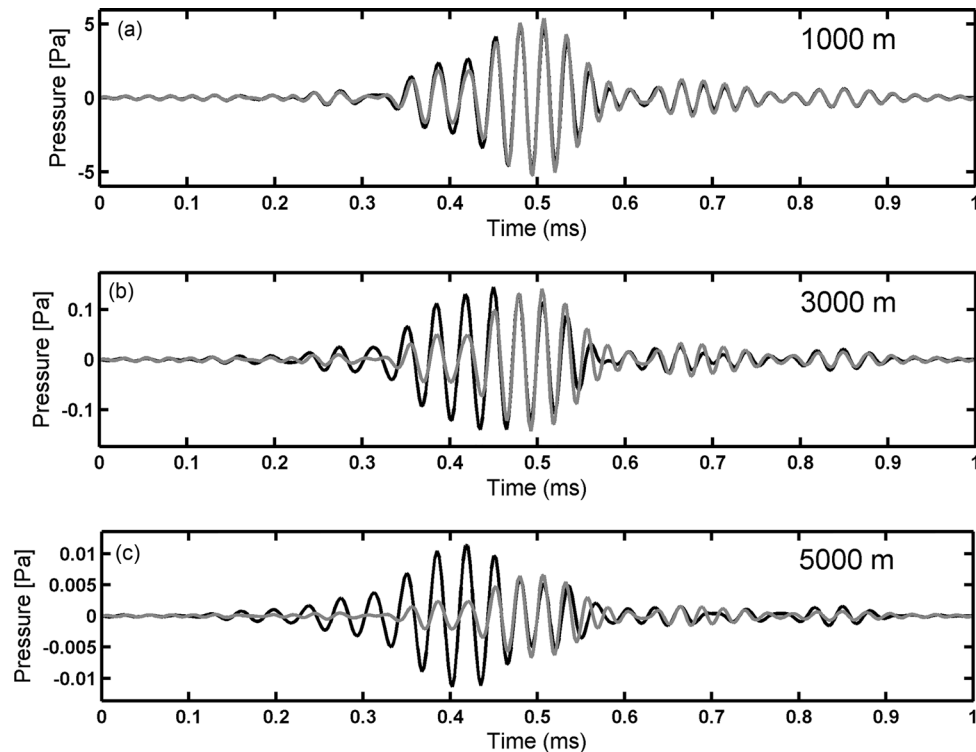


FIG. 2. Example of a Blainville's beaked whale sound pressure time series propagated to different range using the narrowband assumption (gray), and broadband assumption (black) at different source range (top to bottom: 1000, 3000, and 5000 m). The resulting waveform has a different amplitude, due to the overestimation of the propagation loss in the narrowband assumption. The pulse-shape is also clearly different, because the low frequency component is more strongly attenuated in the narrowband approach than in the broadband approach.

For background noise, acoustic measurements were used that were obtained with a towed array (von Benda-Beckmann *et al.*, 2010) at the US Atlantic Undersea Test and Evaluation Center (AUTEC), a U.S. Navy range where Blainville's beaked whales are commonly present and sperm whales sometimes occur (Marques *et al.*, 2009; Ward *et al.*, 2011). Selections were audited by human operators to ensure that no marine mammal vocalizations were present during the time of recording. The Navy acoustic monitoring system provided an independent check to ensure absence of animal vocalizations detected on the array (von Benda-Beckmann *et al.*, 2010). A segment of three hour duration was selected for estimating the false alarm rates from the measured background noise during the AUTEC09 trial using the towed hydrophone array.

From this time series, noise segments of 20 s duration were selected at random for inserting synthetic propagated clicks to determine click detection probability  $p_d$ .  $N_{\text{clicks}}$  clicks were distributed randomly in a selected fragment of background noise, resulting in a timeseries  $x_{S+N}(t)$ , containing the synthesized signals added to the measured background noise.

The detection probability for a single on-axis click as a function of source distance,  $g(r)$  (i.e., the detection function) was modelled for two different detectors: a sub-band energy detector, and a Page-test detector (see Appendix A 2 for a full description). The detection function was obtained by inserting 200 realizations of clicks propagated from distance  $R$  into the selected background noise at random times. The distance  $R$  ranged between 100 and 5000 m for Blainville's beaked

whales, and between 100 and 20 000 m for sperm whales. The detector was run using different fixed detector thresholds.

Realistic detector settings were chosen that corresponded to realistic and practicable false-alarm rates (FAR). A reasonable FAR is typically between 1–10 per hour in practice (e.g., Marques *et al.*, 2009), but may be higher for a single click if no additional classification has yet been carried out (for example, by a human auditing or automatic species classification). To delimit a typical practical range of FAR for a single click detection, we considered two cases: a “low FAR” case with FAR of 1/h, and a “high FAR” case with a FAR of 200/h. The methods used to establish the FAR is described in the Appendixes.

The detection function was estimated for each combination of species, detector, and narrow/broad-band model by fitting a binary generalized additive model (Wood, 2017) to the number of inserted clicks that were successfully detected by the detector. To ensure that the fitted function was monotonic non-increasing, a constrained B-spline basis function was used (Pya and Wood, 2015). The fitted function was used to estimate  $A_{d,\text{eff}_{\text{NB}}}$  and  $A_{d,\text{eff}_{\text{BB}}}$  for each combination of species and detector using Eq. (2). The relative percentage bias was then calculated using Eq. (3).

## C. Using the passive sonar equation to model probability of detecting echolocation clicks

### 1. Passive sonar equation

The passive sonar equation is commonly used to predict detectability of echolocation clicks (Au *et al.*, 2004; Zimmer

*et al.*, 2008; von Benda-Beckmann *et al.*, 2010; Küsel *et al.*, 2011; Küsel *et al.*, 2016; Ward *et al.*, 2011; Ainslie, 2013), and can be a useful tool to model the detection performance of signals at different distances or in different noise conditions. The equation is typically used in the mean square pressure form or in the energy form (Urlick, 1983; Au, 2014).

The energy form of the passive sonar equation for the signal excess (SE) in a specified frequency band can be written in terms of the energy source level (ESL), the propagation loss (EPL), the energy noise level (ENL), and the detection threshold (DT) (Ainslie, 2010)

$$SE = ESL - EPL - ENL - DT. \quad (7)$$

The detection on an omni-directional hydrophone was considered here, for which the processing gain was assumed to be negligible.

Because the click energy is unevenly distributed in both frequency and time (Fig. 1), it is not straightforward to characterize the click in the passive sonar equation. In this study, two different approximations of spectral energy distribution were considered in the passive sonar equation: a top-hat distribution within the frequency band for the detector (Küsel *et al.*, 2011), and within-the-click effective bandwidth (Ainslie, 2013). The same energy source level was adopted as for the simulated waveforms (Fig. 3).

For propagation loss we considered spherical spreading loss and absorption loss. For the sound absorption, two approximations made: a narrowband approach in which the absorption spectrum was approximated by its value at the center frequency of the click (as in Zimmer *et al.*, 2008; Küsel *et al.*, 2011; Ward *et al.*, 2011) and a broadband approach where the absorption was approximated by a linear dependence on frequency within the equivalent bandwidth of the echolocation click (von Benda-Beckmann *et al.*, 2010; Ainslie, 2013).

For the background noise, two cases were considered: (1) the system-noise limited energy noise level of the Delphinus array during the AUTE09 trial, which was measured within the species bandwidth within 1 ms time windows (Table I); (2) a typical wind noise and thermal-limited background spectral noise levels for a sea state of SS = 2, based on Wenz

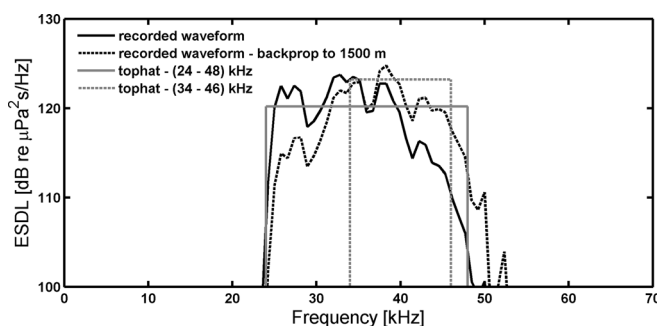


FIG. 3. Comparison of the level of ESDL of the source waveforms, for different assumptions of the energy distribution of an on-axis Blainville's echolocation click. Each curve was normalized to provide an equal broadband energy source level of  $ESL = 164 \text{ dB re } 1 \mu\text{Pa}^2 \text{ m}^2 \text{ s}$ . Note that due to differences in assumed spectral shape, the amplitude of the ESDL differs between model approaches.

(1962) converted to an energy noise level by correcting for the bandwidth (BW) and integration time of 1 ms,  $ENL = NL_{\text{Wenz}} + 10\log_{10}(BW/1 \text{ Hz}) \text{ dB} + 10\log_{10}(0.001 \text{ s}/1 \text{ s}) \text{ dB}$ . This Wenz conversion resulted in ambient noise levels that were similar (within 5 dB) to the ambient noise levels measured on bottom-mounted hydrophones at the AUTE09 range (Ward *et al.*, 2011).

The detection probabilities of an ensemble of clicks as function of signal-to-ratio were calculated by propagating the source waveform to a distance  $R = 1000 \text{ m}$ , and inserting 200 realizations of this click at random times, but keeping the ratio fixed. Here SNR was defined as the difference between the time-integrated squared sound pressure level of the signal and the noise measured within a frequency bandwidth BW around the central frequency  $f_c$  (from Table I). The integration time for determining the SNR was 1 ms. In total this process was repeated for 20 realizations and the number of true detections,  $n_d$  was determined. For each detector threshold, the mean detection probability  $p_d = n_d / n_{\text{click}}$  was computed as a function of SNR (Fig. 4). From these distributions, we derived the detection threshold (DT, not to be confused with the *detector* threshold defined earlier as part of the click detection algorithm), defined as the value of the SNR at which the probability of detection is 50%.

## 2. Determining the relative bias in detection area for other echolocating marine mammal species using the passive sonar equation

To determine the magnitude of the relative bias in estimated population density for echolocating species other than the sperm whale and Blainville's beaked whale, we combined the passive sonar equation approach in Ainslie (2013) with the detection thresholds obtained in the present study. Energy source levels for the different echolocating species for which Ainslie (2013) estimated the relative bias in density estimation as a function of the figure of merit (FOM) were taken from the literature (Table II). For the energy

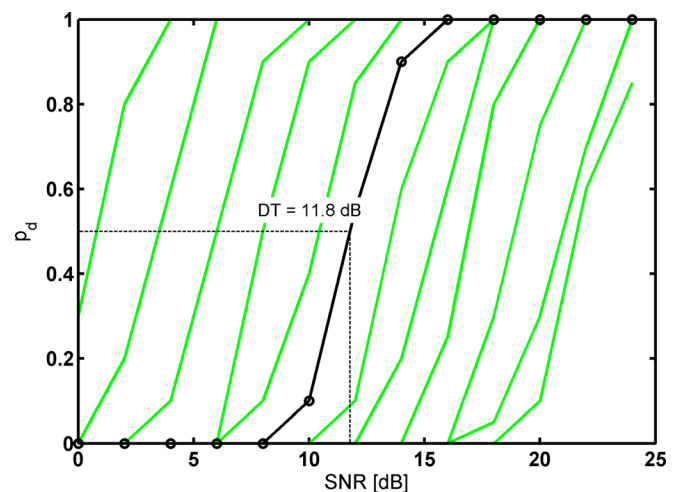


FIG. 4. (Color online) Example of modeled detection probabilities,  $p_d$ , as a function of SNR of Blainville's beaked whale clicks using a simple sub-band energy detector for different detector thresholds. The detection threshold, DT, was determined at the point of 50% detection probability (dashed lines).

TABLE II. Source parameters for echolocation clicks of other marine mammal species (see Ainslie, 2013), and energy source level for these species obtained from the literature (Møhl *et al.*, 1990; Madsen *et al.*, 2004; Au *et al.*, 2004; Wahlberg *et al.*, 2011). Note that the killer whale energy source level was estimated from the peak-to-peak source level from Au *et al.* (2004), by correcting for the pulse shape (Ainslie, 2010) using an assumed pulse duration of 100  $\mu$ s (Au *et al.*, 2004). The next to last two columns list the mean energy noise level ENL within these frequency bands as measured in the AUTECH09 dataset and the last column lists ENL based upon Wenz (1962) for Sea State 2.

Species	$f_m$ [kHz]	$B_{1/2}$ [kHz]	$B_{eq}$ [kHz]	Mean ESL [dB re 1 $\mu$ Pa <sup>2</sup> m <sup>2</sup> s]	Reference	Mean ENL—AUTECH09 [dB re 1 $\mu$ Pa <sup>2</sup> s]	ENL—Wenz (1962) for Sea State 2 [dB re 1 $\mu$ Pa <sup>2</sup> s]
<i>Grampus griseus</i> Risso's dolphin	49.0	27.0	28.7	193	Madsen <i>et al.</i> (2004)	70.1	42.3
<i>Monodon monoceros</i> Narwhal	40.0	20.0	21.3	174	Møhl <i>et al.</i> (1990)	69.3	43.0
<i>Pseudorca crassidens</i> False killer whale	43.0	33.0	35.1	157	Madsen <i>et al.</i> (2004)	71.3	44.2
<i>Orcinus Orca</i> Killer whale	45.0	26.7	28.4	161	Au <i>et al.</i> (2004)	71.3	44.2

form of the passive sonar equation, the FOM is the difference between the energy source level, ESL, and the sum of energy noise level ENL and the detection threshold (ignoring processing gain),

$$\text{FOM} = \text{ESL} - (\text{ENL} + \text{DT}). \quad (8)$$

To assess the order of magnitude of the relative bias in density estimate for species other than Blainville's beaked whale and sperm whales, the detection function was simplified using a cookie-cutter approach (a step function at the DT, i.e., at a 50% detection probability), which was adopted to translate the bias in detection range to a bias in density (Ainslie, 2013). Mean energy noise levels in the frequency band of the animal clicks were measured in the AUTECH09 dataset (Table I and Table II).

### III. RESULTS

To investigate the effect of model assumptions on the modeled detection functions (probability of detecting a single click as a function of animal distance) and resulting relative biases in the click-based animal density estimates, detection thresholds were first established for different species and detectors. The consequence of these detection thresholds for the relative bias on the density estimation was assessed using the passive sonar equation and using the method in which simulated echolocation clicks are injected into background noise. Finally, the passive sonar approach and the simulated data approach were compared.

#### A. Detection thresholds for different detectors and echolocating species using simulated echolocation clicks

To estimate the relative bias on the click-based animal density with the passive sonar model approach, for a fixed false-alarm probability, realistic detection thresholds were established for different species (Fig. 5). The DT so determined was found to depend on the species and background noise statistics considered. A low false-alarm rate of  $\text{FAR} = 1/\text{h}$  was achieved for detection thresholds between 10 and 24 dB. For a high  $\text{FAR} = 200/\text{h}$ , detection thresholds ranged between 1 and 12 dB. The overlap in DT yielding such different FAR values between the two detectors for beaked whale clicks illustrates how the DT required for a

specific FAR can be strongly detector dependent. Detection thresholds that allowed for low FAR of  $1/\text{h}$  were between 10 and 20 dB (Fig. 3).

#### B. Effect of the narrowband approximation of absorption on the detection function—passive sonar equation

A measurement of the DT for a specific FAR allowed us to quantify the relative bias on the effective detection area obtained by Ainslie (2013) (Fig. 5). To provide a range of relative bias predicted by the passive sonar equation for different species listed in Tables I and II, two background noise limits were defined: (1) dominated by electronic interference (system noise), with a high detection threshold (leading to low detection ranges); (2) dominated by noise generated by the sea surface (for a sea state = 2), with a low detection threshold (indicated by the circles in Fig. 6).

The predicted relative density bias due to the narrowband assumption for the system noise case ranged from 5% to 25% for most species, with 42% for Risso's dolphin. For the ambient noise case, the relative bias increased to values between 20% (*Mesoplodon*) and 115% (*Pseudorca*). The predicted range of relative bias was lowest for beaked

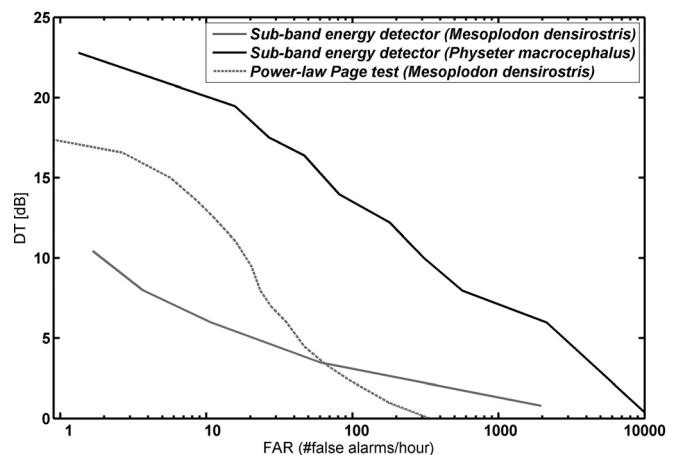


FIG. 5. Detection threshold vs false-alarm rate (FAR), obtained by varying the detector thresholds, for detection of *Mesoplodon densirostris* and *Physeter macrocephalus* echolocation clicks. The line styles indicate different detectors used (dashed, energy sub-band detector; solid: power law Page test detector). Note how a wide range of DT can be obtained for a given FAR, depending on the type of detector used, and species considered.

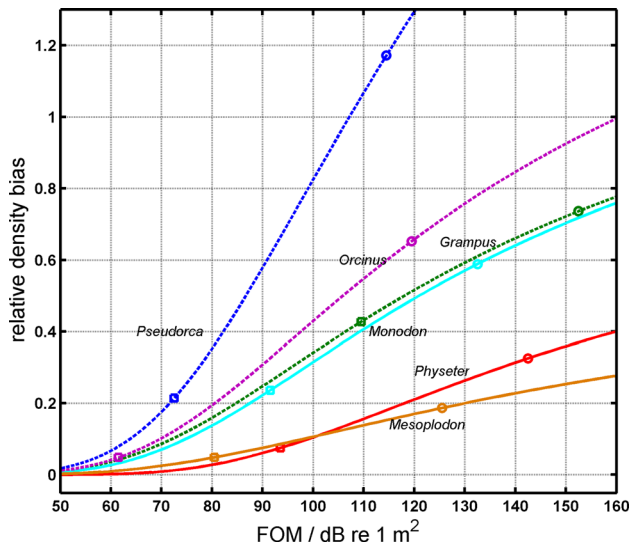


FIG. 6. (Color online) Modeled relative bias in density estimate between narrowband and broadband approximation as a function of figure of merit (FOM). The squares are for a relatively high DT = 20 dB, and the detection in noise conditions of the towed array at the AUTECH09 trial, and circles indicate the detection performance when the system is detecting in noise limited by wind and thermal noise conditions, for a sea state of SS = 2 (Wenz, 1962), using a low detection threshold (DT = 5 dB). These symbols indicate a realistic range of FOM that can be expected for different species, and the relative importance of frequency dependent absorption in the detection model.

whales, and sperm whales. Figure 6 illustrated that the FOM was strongly species dependent, and a fixed FOM was not a good predictor of the relative density bias.

### C. Effect of the narrowband approximation of absorption on the detection function—simulated echolocation clicks

Detection functions for Blainville's beaked whale and sperm whale echolocation clicks were estimated using the broadband and narrowband approximations of the absorption using the simulations with the recorded waveforms (Fig. 7). For the combination of background noise and detector settings used in the sub-band energy detector, effective detection radii ( $R_{d,eff}$ ) for on-axis clicks were predicted to range from 1.1 to 1.8 km (broadband) and 1.0 to 1.6 km (narrowband) (7%–10% error) for the beaked whales, with maximum detection ranges up to 2.3 km. A high probability of detection ( $p_d$  close to unity) was predicted for on-axis clicks at distances out to 0.8–1.5 km (broadband), and 0.7–1.3 (narrowband). The resulting detection function for sperm whales showed much larger detection distances, with  $R_{d,eff}$  = 6.5–13.8 km (broadband) and 4.9–8.9 km (narrowband), with a larger difference between the NB and BB approximations (25%–35%). The relative bias in density estimate was 11%–21% for Blainville's beaked whales, and 76%–138% for sperm whales (Table III).

The sub-band energy detector and the power-law Page test detector predicted different effective detection radii for a similar FAR. For a low FAR (high detection threshold), the broadband effective detection radius predicted for the

power-law Page test detector was smaller than for the sub-band energy detector (0.84 vs 1.03 km), whereas at high FAR (low detection threshold), the power-law Page test predicted slightly higher broadband effective detection radius (1.86 vs 1.80 km).

### D. Effect of assumptions on click spectral energy distribution on the detection function for Blainville's beaked whales

The effect of differences in assumed click energy distribution was investigated by assuming the same broadband energy flux of the Blainville's beaked whale click and varying the spectral energy distribution. The noise conditions were kept fixed, and we used the sub-band energy detector with a low DT. Different approximations for the spectral energy distribution using the simulated waveform and passive sonar equation resulted in detection functions with effective detection areas ranging from 8.8 to 17.4 km<sup>2</sup> (Fig. 8 and Table IV). The simulation in which the on-axis click was backpropagated predicted the smallest effective detection radius. When the click energy distribution was approximated by a top-hat with a relatively large bandwidth, the passive sonar equation resulted in a relatively large effective detection radius. This was a consequence of the source click waveform having relatively more energy at lower frequencies, which were less strongly attenuated and thus the clicks were detectable at larger distances.

## IV. DISCUSSION

Detection functions of echolocation clicks are often based on modeling the detectability of single clicks. These predictions often rely on modeling using the passive sonar equation, which may be non-trivial for some complex detectors and which requires assumptions to be made about the beam pattern and effective bandwidth of the echolocating clicks transmitted by the animals. This study compared methods to model the detectability of broadband clicks, using a new approach where recorded echolocation clicks were inserted into measured background noise and by applying different detectors in order to assess the effect of common approximations made when modeling the propagation of broadband echolocating clicks using the passive sonar equation.

### A. Effect of the approximations in the sonar equation on modeled detection functions

This study demonstrated that ignoring the frequency dependence of absorption in the propagation modelling resulted in an underestimation of the effective detection radius (or area) for some marine mammal species. The models also predicted different detection functions for different common approximations made depending on how the source spectral energy of echolocation clicks is distributed. The estimated errors in effective detection radius and resulting relative bias in density estimates were computed using different approaches (using waveforms injected in background noise data, and the passive sonar equation



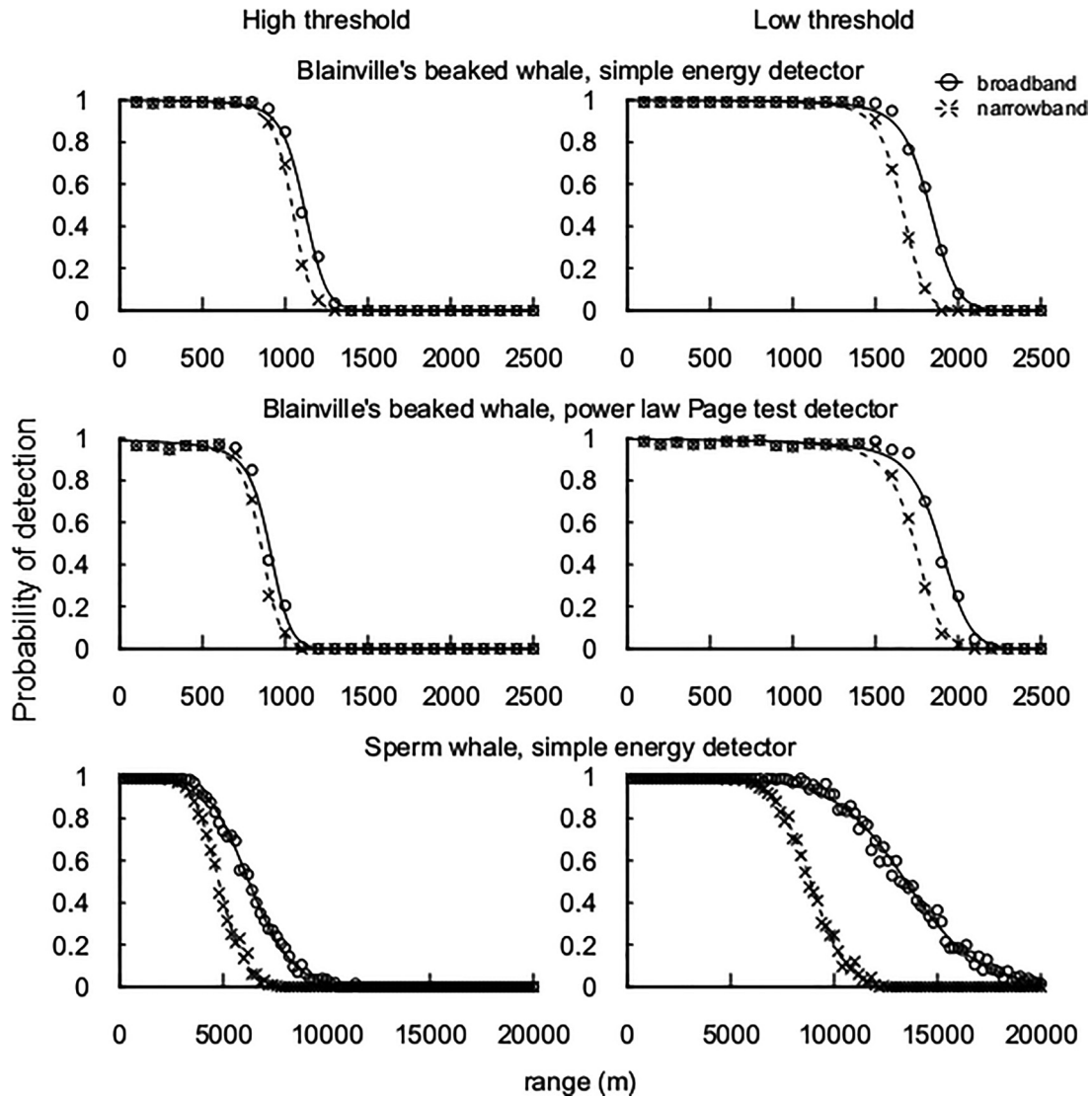


FIG. 7. Modeled detection functions for on-axis clicks, using simulated waveforms, propagated with a broadband (circles) and narrowband (crosses) approximations for absorption. Top: Detection function of a single on-axis *Mesoplodon densirostris* (Blainville's beaked whale) click using a simple energy detector (with DT = 7 dB) corresponding to a FAR = 1/h, and DT = 3 dB, corresponding to a FAR ~ 200/h). Middle: Detection of a Blainville's beaked whale click using a power law Page test detector (with DT = 12 dB) corresponding to a FAR = 1/h, and DT = 2 dB, corresponding to a FAR ~ 200/h). Bottom: Detection function of a single on-axis *Physeter microcephalus* (sperm whale) click using a simple energy detector (with DT = 23 dB, corresponding to a FAR = 1/h, and with DT = 11 dB, corresponding to a FAR = 200/h).

with top-hat approximations of the click spectrum) for similar background noise and detection thresholds, allowing a direct comparison of the two methods (Fig. 8 and Table IV).

The top-hat passive sonar approach for a center frequency of 36 kHz and bandwidth of 24 kHz (Fig. 3) resulted in the largest effective detection area, which was 68% larger than the effective detection area modeled with the recorded

TABLE III. Effective detection areas ( $A_{d,eff}$ ) and relative density biases obtained from a fit to the modeled detection function, for Blainville's beaked whales and sperm whales, for different detectors and detection thresholds used. Results are shown for the broadband assumption and the narrowband assumption. The narrowband approximation of the absorption resulted in a relative bias in the estimated density from Eq. (3) (last column). Detection thresholds were chosen to provide the same false-alarm rate of FAR = 1/h (high DT), and FAR = 200/h (low DT), in a system-noise limited background noise (AUTEC09 data).

	Species	Detector type	FAR	$A_{d,eff}$ [km <sup>2</sup> ] (BB)	$A_{d,eff}$ [km <sup>2</sup> ] (NB)	Relative density bias
1	Blainville	Energy sum	Low (1/h)	3.9	3.4	0.15
2	Blainville	Energy sum	High (200/h)	10.3	8.5	0.21
3	Blainville	Power law	Low (1/h)	2.5	2.2	0.12
4	Blainville	Power law	High (200/h)	10.9	9.2	0.18
5	Sperm	Energy sum	Low (1/h)	134.0	76.1	0.76
6	Sperm	Energy sum	High (200/h)	596.3	250.8	1.38

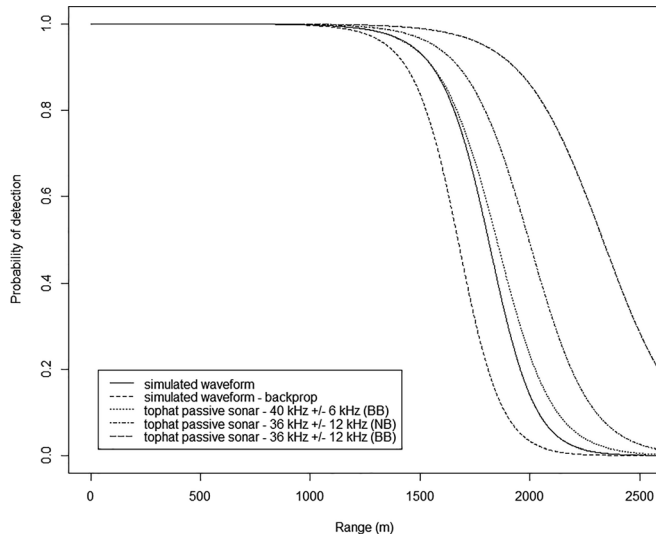


FIG. 8. Comparison of fitted detection functions to data simulated using different assumptions on the source waveform spectral energy distribution for a Blainville's beaked whale echolocation click. Scenarios are for the simple subband energy detector with a high FAR = 200/h, in a system-noise limited background noise (AUTC09 data), and broadband approximations for the sound absorption.

waveform (Fig. 8). The top-hat passive sonar model with a constant absorption across all frequencies (corresponding to the Küsel *et al.*, 2011 study) resulted in a somewhat smaller effective detection area than for the broadband approximation with the same bandwidth. However, the effective detection area was 23% higher than that predicted with the recorded waveform. The top-hat passive sonar approach with the energy distributed in the center frequency of 40 kHz and an effective bandwidth of 12 kHz (as in Ainslie, 2013) led to a slightly higher (5%) effective detection area than for the recorded waveform.

The on-axis beaked whale echolocation click shown in Fig. 1 was recorded some distance away from the animal, thus the click spectrum may be biased towards lower frequencies due to frequency-dependent absorption. Reliable distances to the animal emitting this click could not be determined, but based on detections on hydrophones detecting beaked whales at the AUTC09 range during the AUTC09 trial (von Benda-Beckmann *et al.*, 2010), we estimated the distance of the animal group to be between 1000 and 2000 m from the receiver. To assess the effect of using a received signal instead of the actual source signal, the received waveform depicted in Fig. 1 was back-computed to the approximate source location (assumed to be 1500 m), resulting in a

click waveform that has a higher frequency content (see Fig. 3). The peak frequency and bandwidth of the back-propagated waveform were also in agreement with literature values for Blainville's beaked whale source clicks (Johnson *et al.*, 2004). Back-propagation of the click to its approximate location (~1500 m) resulted in a somewhat smaller effective detection area than for the recorded waveform (Table IV).

The relative bias as a result of neglecting the broadband nature of clicks depended upon the species studied and the detector used (Fig. 6). Different detectors led to somewhat different detection functions for the same FAR. For the different detectors adopted in this study, a range of DT between 5 and 20 dB was obtained. Using this range of detection thresholds with typical background noise levels, the corresponding range of FOM modeled in Ainslie (2013) suggested that the smallest relative bias for the species shown in Fig. 6 was for beaked whales, with an increase for species with very broadband clicks, such as the false killer whale, and Risso's dolphins. Note that these estimates in relative bias were based on the simplifying assumption that the detection function could be considered as a step-function at the detection threshold. In order to compute the true bias in effective detection range, the change of  $p_d$  as a function SNR is required, which was done in this study for beaked whales and sperm whales. For the other species, these values should therefore be taken as an order of magnitude estimate.

For the sperm whale, the passive sonar equation predictions of the relative density bias due to the narrowband assumption of sound absorption were much smaller than those obtained with the model using the recorded waveform (8%–35%, compared to 76%–137%). Ainslie (2013) considered a top-hat approximation of the click, using the equivalent bandwidth. The sperm whale click presented here had a somewhat lower peak frequency, and hence more energy content below the -3 dB band, than the one used by Ainslie (2013), which was based on Möhl *et al.* (2003). Consequently, the click attenuated less and was detectable to larger distances, increasing the detection range and thus increasing the resulting relative bias of the density estimate. In order to disentangle the two effects, the source waveform in Fig. 1 needs to be back-propagated to obtain a more representative on-axis source click spectrum of sperm whales (as, for example, described in Möhl *et al.*, 2003). This was not attempted for the sperm whale click in this study, since the actual source positions of the vocalizing whale was not available.

TABLE IV. Comparison of effective detection area ( $A_{d,eff}$ ) obtained using different assumptions about the source waveform spectral energy distribution of an on-axis Blainville's beaked whale echolocation click (see Fig. 4). The third column lists the type of assumptions made for the frequency dependent absorption. Scenarios are for the simple subband energy detector with a high FAR = 200/h, in a system-noise limited background noise (AUTC09 data).

	Model type	Absorption loss assumption	$A_{d,eff}$ [km <sup>2</sup> ]
1	Simulated waveform: backcomputed 1500 m	Frequency dependent	8.8
2	Simulated waveform	Frequency dependent	10.3
3	Passive sonar top-hat: (34–46) kHz	Linear dependence in signal bandwidth	10.8
4	Passive sonar top-hat: (24–48) kHz	Linear dependence on frequency in signal bandwidth	17.4
5	Passive sonar top-hat: (24–48) kHz	Constant at 38.5 kHz	12.7

From Table IV it can be deduced that even though the narrowband approximation of the absorption tended to result in smaller effective detection areas, approximating the spectral energy distribution by a top-hat function with a wide frequency range still resulted in a larger predicted effective detection areas than those predicted with the propagated waveform. The range in predicted effective detection areas of 8.8 to 17.8 km<sup>2</sup> was based on the same broadband energy source level and detector settings. The range of effective detection areas was solely driven by the assumptions about the assumed bandwidth and how the click spectral energy density was approximated. This highlights the importance of accounting for the broadband nature of echolocation clicks in model predictions of detection functions.

## B. Other considerations for estimating the effective detection area

In this study, *single-click* detection probabilities were considered for estimating density of echolocating species. Since many odontocete species often produce sequences of clicks, their detection can be significantly improved by considering multiple clicks (Zimmer *et al.*, 2008; von Benda-Beckmann *et al.*, 2010). Group presence can be used, for instance, to detect start and end of dives of some species, such as beaked whales, which can also be used as a cue to establish animal densities (Moretti *et al.*, 2010). Detecting the presence of multiple consecutive clicks with a species-specific range of inter-click intervals, is a common practice to reduce the number of false alarms, thus allowing for a lower detection threshold. In principle, the effect of reducing the FAR by classification can be modeled explicitly by inserting sequences of clicks separated in time by observed inter-click-intervals. Here the effect of classification of multiple clicks on the detection ranges is illustrated using a low DT, which produced a fairly high FAR when single clicks were considered (Fig. 4), but in practice led to acceptable false alarm rates, when group presence within 2 to 5 min intervals was considered (von Benda-Beckmann *et al.*, 2010; Hildebrand *et al.*, 2015). Lower detection thresholds allow for detection at larger distances, and therefore can increase the potential relative bias (Fig. 6).

An important question is to what extent the results are robust to the waveform adopted for the two species in this study. In practice, received waveforms are strongly aspect dependent (Møhl *et al.*, 2000; Zimmer *et al.*, 2005; Shaffer *et al.*, 2013). Here, only on-axis clicks were considered for both beaked whales and sperm whales. Average detection functions for an ensemble of on-axis and off-axis clicks can be estimated using model assumptions about animal diving behavior (Zimmer *et al.*, 2008; von Benda-Beckmann *et al.*, 2010; Küsel *et al.*, 2011; Hildebrand *et al.*, 2015), or can be empirically derived when animal locations can be measured (by using tags to track the animals movement, or by passive acoustic tracking, e.g., Marques *et al.*, 2009; Ward *et al.*, 2011).

The sensitivity of the detection performance to the on-axis click waveform was investigated by propagating the source waveform of the beaked whale click to different

distances, and estimating the detection thresholds using the sub-band energy detector. The detection threshold was found to be insensitive to the adopted waveform, with a standard deviation of 1 dB, for this particular detector. Off-axis clicks tend to be more narrow band than on-axis clicks (e.g., Wahlberg *et al.*, 2011; Shaffer *et al.*, 2013), and therefore may lead to lower relative bias than on-axis clicks. However, because on-axis clicks have a higher source level, they are more likely to be over-represented in detections at large distances, and therefore would dominate the detection function for an ensemble of clicks. It is therefore expected that the errors estimated here would be similar to those as estimated for detection of click trains.

A mean energy source level was assumed, which was based on literature values. However, energy source level varies considerably within and between individuals (Shaffer *et al.*, 2013). In the modelling approach for simulating acoustic data, the variability in click source level was considered by including a Gaussian distribution with a standard deviation of ~1.5–3.5 dB, following the approach in Küsel *et al.* (2011). However, uncertainty on that mean itself was not considered, which can lead to an artificially low variance in the estimated density from the Monte-Carlo approach. When applying modeled detection functions to density estimation, variability in mean source level needs to be accounted for.

The estimates given in this study were for a situation in which the background noise suffered from electronic interference (see Sec. II). Figure 6 provides an indication how these estimates may change when the detection is performed in ambient noise limited conditions resulting from wind-generated noise, and suggests that the relative bias may be a factor 2 to 12 higher for conditions limited by ambient noise.

In many cases considered here, wind-generated and thermal noise dominate the background. However, ship noise may dominate for lower frequencies in the case of towed arrays, or during close passage of ships near bottom-mounted systems, which would need to be accounted for when estimating the detection functions. We recognize that a more rigorous sampling procedure would be necessary to select representative samples of noise in a real-world detector characterization scenario, but for the purposes of this paper, which is to compare different narrowband and broadband approximations to quantify the detector performance, the above approach was deemed sufficient. Estimation of the bias for a specific site would require noise data from that site.

The simplified spherical spreading model for propagating the echolocation clicks was used in this study to enable straightforward comparison in the relative bias that stems from neglecting broadband absorption. For sources that are relatively nearby (<2 km) within deep water, this is a reasonable approximation. For sperm whale echolocation clicks, which can be detected at large distances, spherical spreading may underestimate the true detection distance, for instance, due to presence of a sound channel, or overestimation in the case when the receiver is in the shadow zone (e.g., Madsen *et al.*, 2006). This biased detection distance would lead to a larger relative bias than predicted in Fig. 5. Models that capture the spreading characteristics of sound suitable for the acoustic environment can be adopted when applying modeled

detection functions in real-life situations (e.g., Küsel *et al.*, 2011; Küsel *et al.*, 2016; Ward *et al.*, 2011, Helble *et al.*, 2013). Model predictions could be further validated by playing back echolocation signals at different distances, and recording these in realistic environments.

The focus of this study was on how uncertainties in the modelled effective detection area,  $A_{d,eff}$  affect the density estimation. Other parameters from Eq. (1) used in the density estimation, such as the cue rate, distribution and orientation of vocalizing animals, also introduce significant variability and uncertainty, which can have large effect on the density estimation (e.g., see Marques *et al.*, 2013, and references therein). Model-based approaches for deriving the effective detection area will always involve uncertainties (due to natural variability in propagation conditions, noise, and animal vocal behavior), and the uncertainties should be seen in the context of uncertainties in the other parameters in Eq. (1). A comparison of the uncertainties in these different parameters is beyond the scope of the current study; however we emphasize that unnecessary biases resulting from simplifications made to model the effective detection area should be avoided, especially if they are significant compared to the uncertainty resulting from other parameters, such as animal click-rate, group size, orientation.

## V. CONCLUSION

This paper compares different model approaches for predicting detection functions of broadband marine mammal echolocation clicks that are used for estimating marine mammal densities. A method of injecting simulated echolocation clicks into recorded acoustic noise data was compared to model estimates using the passive sonar equation based on different approximations of broadband clicks found in the literature. Simulated waveforms were obtained from recorded on-axis clicks that were propagated to the receiver distance to predict effective detection areas under different noise conditions and using different detectors.

The model comparison made in this study demonstrated how differences in model approaches can result in significant differences in predictions of the effective detection area. By comparing model predictions with different assumptions about the spectral energy distribution of a Blainville's beaked whale echolocation click, it was found that the predicted effective detection area could differ by a factor of 2, even when exactly the same energy source level and detector settings were used.

Both model approaches were used to demonstrate how commonly used simplifying narrowband approximations to the propagation of broadband clicks can significantly underestimate the detection function, resulting in a potential for bias in the click-based density estimates. The relative density bias as a consequence of these simplifying assumptions could be quantified for different other odontocete species, and ranged between 2% to more than 100% for some species, depending on the species, detector settings and noise conditions considered.

We recommend caution in applying narrowband approximations when modeling broadband echolocation clicks, as these have the potential to bias population density estimates.

The alternative methods proposed here and in other studies (e.g., Helble *et al.*, 2013; Küsel *et al.*, 2016), which rely on the availability of high quality source waveforms, can capture the broadband nature of echolocation clicks by propagating the source waveforms to the receiver position and injecting them in noise data.

## ACKNOWLEDGMENTS

This project was funded by the U.S. Office of Naval Research (ONR Grant No. N00014-14-1-0409). We thank ONR program officer Dr. Michael Weise for his support. We further thank Walter Zimmer for making the sperm whale echolocation click recordings available to be used in this study. P.L.T. acknowledges funding received from the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland). MASTS is funded by the Scottish Funding Council (Grant No. HR09011) and contributing institutions. We thank two anonymous reviewers for their constructive comments, which helped to improve this manuscript.

## APPENDIX A: CLICK DETECTORS

### 1. Sub-band energy summation detector

The sub-band energy detector was implemented by computing the time-integrated squared sound pressure  $E_{sum}$  within the frequency band from  $f_{min}$  to  $f_{max}$  over consecutive time windows of  $N_{FFT}$  elements. The  $N_{FFT}$  was set to 2048 samples, which for a sampling rate of 400 kHz corresponded to a time window of 5.22 ms. A detection was considered made whenever the energy exceeded the detector threshold (Abraham, 2010), i.e.,  $E_{sum} \geq D_{det}$ .

### 2. Power-law energy Page-test detector

The power-law Page test detector is described in detail in van IJsselmuide *et al.* (2004). First the background noise was estimated, which was used for normalizing the power spectrum of the signal. Next, a power-law integrator was applied to the normalized spectrum. Based on the integrated normalized spectrum, a Page test (Page, 1954) was used to determine whether a transient signal was detected. The power-law Page detector was run over small elements of  $N_{FFT}$  samples and detected onset and offset of transient signals using various settings, which accounted for the variability in background noise. Here detections were characterized by a start time  $T_{start}$  and end time  $T_{end}$ , and could cover multiple consecutive elements of size  $N_{FFT}$ . A true detection was defined as a detection where the time of the click (center of inserted waveform) was within the start and end times ( $T_{start}$  and  $T_{end}$ ) of the time of detection.

## APPENDIX B: MEASURING FALSE-ALARM RATES

The false-alarm rate (FAR) of a detector was estimated by varying the detector thresholds  $D_{det}$  and running the detector over a 3 h segment of background noise to determine the number of false detections,  $N_{fa}$ . The probability of false alarms is defined as  $p_{fa} = (n_{det}/n_{do})$ , with  $n_{det}$  as the



number of false detections and  $n_{do}$  the number of detection opportunities. In the case of the simple energy detector, the number of detection opportunities in one hour was  $n_{do} = (3600\text{ s}) f_s / N_{FFT}$ . The probability of false alarms for the simple energy detector corresponding to a FAR = 1/h to 10/h was  $p_{fa} = 1.6 \times 10^{-7}$  to  $1.6 \times 10^{-6}$ .

A reasonable FAR is typically between 1–10 per hour in practice (e.g., Marques *et al.*, 2009), but may be higher for a single click if no additional classification has yet been carried out (for example, by a human auditing, or automatic species classification). To delimit a typical practical range of FAR for a single click detection, we considered two cases: a “low FAR” case with FAR of 1/h, and a “high FAR” case with a FAR of 200/h.

Because the power-law Page test depends on an adaptive history of background noise estimates (Page, 1954; van IJsselmuiden *et al.*, 2004), the detection probabilities for individual bins of size  $N_{FFT}$  were not strictly independent of each other, resulting in a non-trivial relationship between the number of false-alarms and the false-alarm probability. However, since the detector settings were set to detect single clicks with a duration smaller than 0.001 s (4000 bins), these bins were treated as if they were independent. For the low detector threshold, the complete time series could be considered as one continuous detection. In such cases, consecutive bins with false detections were considered as a single false-alarm. Therefore, probability of false-alarm was estimated by the ratio of the total duration covered by the transients to the total duration of background noise (e.g., 3 h),  $p_{fa} = [1/(3 \times 3600) \text{ s}] \sum_i T_i$ .

Abraham, D. A. (2010). “Detection-threshold approximation for non-Gaussian backgrounds,” *IEEE J. Ocean. Eng.* **35**(2), 335–365.

Ainslie, M. A. (2010). *Principles of Sonar Performance Modeling* (Springer Verlag, Berlin).

Ainslie, M. A. (2013). “Neglect of bandwidth of Odontocetes echo location clicks biases propagation loss and single hydrophone population estimates,” *J. Acoust. Soc. Am.* **134**, 3506–3512.

Au, W. (2014). “Dolphin biosonar target detection in noise: Wrap up of a past experiment (L),” *J. Acoust. Soc. Am.* **136**(1), 9–12.

Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer-Verlag, New York).

Au, W. W. L., Ford, K. B., Horne, J. K., and Newman Allman, K. A. (2004). “Echolocation signals of free-ranging killer whales (*Orcinus orca*) and modeling of foraging for chinook salmon (*Oncorhynchus tshawytscha*),” *J. Acoust. Soc. Am.* **115**(2), 901–909.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2001). *Introduction to Distance Sampling — Estimating Abundance of Biological Populations* (Oxford University Press, Oxford).

Frasier, K. E., Wiggins, S. M., Harris, D., Marques, T. A., Len Thomas, L., and Hildebrand, J. A. (2016). “Delphinid echolocation click detection probability on near-seafloor sensors,” *J. Acoust. Soc. Am.* **140**(3), 1918–1930.

Gassmann, M., Wiggins, S. M., and Hildebrand, J. A. (2015). “Three-dimensional tracking of Cuvier’s beaked whales’ echolocation sounds using nested hydrophone arrays,” *J. Acoust. Soc. Am.* **138**(4), 2483–2494.

Gillespie, D., and Leaper, R. (1996). “Detection of sperm whale *Physeter macrocephalus* clicks and discrimination of individual vocalizations,” *European Cetacean Research. Abstracts*, pp. 10, 87–91.

Harris, D. (2012). “Estimating whale abundance using sparse hydrophone arrays,” Ph.D. thesis, Statistics and Biology, University of St. Andrews, St. Andrews, Scotland, UK.

Helble, T. A., D’Spain, G. L., Hildebrand, J. A., Campbell, G. S., Campbell, R. L., and Heaney, K. D. (2013). “Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones,” *J. Acoust. Soc. Am.* **134**, 2556–2570.

Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merckens, K. P., Wiggins, S. M., McDonald, M. A., Garrison, L. P., Harris, D., Marques, T. A., and Thomas, L. (2015). “Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico,” *Sci. Rep.* **5**, 16343.

ISO (2017). *ISO 18405 Underwater Acoustics—Terminology* (International Organization for Standardization, Geneva, 2017).

Johnson, M., Madsen, P. T., Zimmer, W. M. X., Aguilar de Soto, N., and Tyack, P. (2004). “Beaked whales echolocate on prey,” *Proc. R. Soc. London B* **271**, S383–S386.

Kandia, V., and Stylianou, Y. (2005). “Detection of sperm whale clicks based on the Teager Kaiser energy operator,” *Appl. Acoust.* **67**, 1144–1163.

Kimura, S., Akamatsu, T., Li, S., Dong, S., Dong, L., Wang, K., Wang, D., and Arai, N. (2010). “Density estimation of Yangtze finless porpoises using passive acoustic sensors and automated click train detection,” *J. Acoust. Soc. Am.* **128**, 1435–1445.

Klinck, H., and Mellinger, D. K. (2011). “The energy ratio mapping algorithm: A tool to improve the energy-based detection of odontocete echolocation clicks,” *J. Acoust. Soc. Am.* **129**, 1807–1812.

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D., and Ward, J. (2011). “Cetacean population density estimation from single fixed sensors using passive acoustics,” *J. Acoust. Soc. Am.* **129**, 3610–3622.

Küsel, E. T., Siderius, M., and Mellinger, D. K. (2016). “Single-sensor, cue-counting population density estimation: Average probability of detection of broadband clicks,” *J. Acoust. Soc. Am.* **140**, 1894–1903.

Madsen, P. T., Johnson, M., Miller, P. J. O., Aguilar Soto, N., Lynch, J., and Tyack, P. (2006). “Quantitative measures of air-gun pulses recorded on sperm whales (*Physeter macrocephalus*) using acoustic tags during controlled exposure experiments,” *J. Acoust. Soc. Am.* **120**(4), 2366–2379.

Madsen, P. T., Kerr, I., and Payne, R. (2004). “Echolocation clicks of two free-ranging, oceanic delphinids with different food preferences: False killer whales *Pseudorca crassidens* and Risso’s dolphins *Grampus griseus*,” *J. Exp. Biol.* **207**, 1811–1823.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (2013). “Estimating animal population density using passive acoustics,” *Biol. Rev.* **88**, 287–309.

Marques, T. A., Thomas, L., Munger, L., Wiggins, S., and Hildebrand, J. A. (2011). “Estimating North Pacific right whale (*Eubalaena japonica*) density using passive acoustic cue counting,” *Endang. Spec. Res.* **13**, 163–172.

Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (2009). “Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville’s beaked whales,” *J. Acoust. Soc. Am.* **125**(4), 1982–1994.

McDonald, M., and Fox, C. (1999). “Passive acoustic methods applied to fin whale population density estimation,” *J. Acoust. Soc. Am.* **105**, 2643–2651.

Mellinger, D. K. (2001). “ISHMAEL 1.0 User’s Guide,” NOAA Technical Memorandum OAR PMEL-120, available from NOAA/PMEL, 7600 Sand Point Way, NE, Seattle, WA 98115-6349.

Möhl, B., Surlykke, A., and Miller, L. A. (1990). “High intensity narwhal clicks,” in *Sensory Abilities of Cetaceans*, edited by J. Thomas and R. Kastelein (Plenum, New York), pp. 295–304.

Möhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (2003). “The monopulsed nature of sperm whale clicks,” *J. Acoust. Soc. Am.* **114**, 1143–1154.

Möhl, B., Wahlberg, M., Madsen, P. T., Miller, L. A., and Surlykke, A. (2000). “Sperm whale clicks: Directionality and source level revisited,” *J. Acoust. Soc. Am.* **107**, 638–648.

Moretti, D., Marques, T. A., Thomas, L., DiMarzio, N., Dilley, A., Morrissey, R., McCarthy, R., Ward, J., and Jarvis, S. (2010). “A dive counting density estimation method for Blainville’s beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to a Mid-Frequency Active (MFA) sonar operation,” *Appl. Acoust.* **71**(11), 1036–1042.

Page, E. S. (1954). “Continuous inspection schemes,” *Biometrika* **41**(1/2), 100–115.

Pyra, N., and Wood, S. N. (2015). “Shape constrained additive models,” *Statistics and computing* **25**(3), 543–559.

Shaffer, J. W., Moretti, D., Jarvis, S., Tyack, P., and Johnson, M. (2013). “Effective beam pattern of the Blainville’s beaked whale (*Mesoplodon densirostris*) and implications for passive acoustic monitoring,” *J. Acoust. Soc. Am.* **133**, 1770–1784.

- Urick, R. J. (1983). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York).
- van IJsselmuide, S. P., and Beerens, S. P. (2004). "Detection and classification of marine mammals using an LFAS array," *Can. Acoust.* **32**, 93–106.
- von Benda-Beckmann, A. M., Beerens, S. P., and van IJsselmuide, S. P. (2013). "Effect of towed array stability on instantaneous localization of marine mammals," *J. Acoust. Soc. Am.* **134**(3), 2409–2417.
- von Benda-Beckmann, A. M., Lam, F. P. A., Moretti, D. J., Fulkerson, K., Ainslie, M. A., van IJsselmuide, S. P., Theriault, J., and Beerens, S. P. (2010). "Detection of Blainville's beaked whales with towed arrays," *Appl. Acoust.* **71**, 1027–1035.
- Wahlberg, M., Jensen, F. H., Aguilar Soto, N., Beedholm, K., Bejder, L., Oliveira, C., Rasmussen, M., Villadsgaard, A., and Madsen, P. T. (2011). "Source parameters of echolocation clicks from wild bottlenose dolphins (*Tursiops aduncus* and *Tursiops truncatus*)," *J. Acoust. Soc. Am.* **130**(4), 2263–2274.
- Ward, J., Jarvis, S., Moretti, D., Morrissey, R., DiMarzio, N., Johnson, M., Tyack, P., Thomas, L., and Marques, T. (2011). "Beaked whale (*Mesoplodon densirostris*) passive acoustic detection in increasing ambient noise," *J. Acoust. Soc. Am.* **129**, 662–669.
- Ward, J., Morrissey, R., Moretti, D., DiMarzio, N., Jarvis, S., Johnson, M., Tyack, P., and White, C. (2008). "Passive acoustic detection and localization of *Mesoplodon densirostris* (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a Digital Tag (DTAG) recording," *Can. Acoust.* **1**, 60–66.
- Watkins, W. A., and Schevill, W. E. (1972). "Sound source location by arrival-times on a non-rigid three-dimensional hydrophone array," *Deep Sea Res. Oceanogr. Abstr.* **10**, 691–692.
- Wenz, G. M. (1962). "Acoustic ambient noise in the ocean: Spectra and sources," *J. Acoust. Soc. Am.* **34**, 1936–1956.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*, 2nd ed. (Chapman and Hall/CRC, New York).
- Yack, T. M., Barlow, J., Roch, M. A., Klinck, H., Martin, S., Mellinger, D. K., and Gillespie, D. (2010). "Comparison of beaked whale detection algorithms," *Appl. Acoust.* **71**(11), 1043–1049.
- Zimmer, W., Harwood, J., Tyack, P., Johnson, M., and Madsen, P. (2008). "Passive acoustic detection of deep-diving beaked whales," *J. Acoust. Soc. Am.* **124**, 2823–2832.
- Zimmer, W. M. X. (2011). *Passive Acoustic Monitoring of Cetaceans* (Cambridge University Press, New York).
- Zimmer, W. M. X. (2013). "Range estimation of cetaceans with compact volumetric arrays," *J. Acoust. Soc. Am.* **134**(3), 2610–2618.
- Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (2005). "Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis," *J. Acoust. Soc. Am.* **117**, 1473–1485.